OPTIMIZING SCHRÖDINGER FUNCTIONALS USING SOBOLEV GRADIENTS: APPLICATIONS TO QUANTUM MECHANICS AND NONLINEAR OPTICS

JUAN JOSÉ GARCíA-RIPOLL * AND VíCTOR M. PÉREZ-GARCíA †

Abstract. In this paper we study the application of the Sobolev gradients technique to the problem of minimizing several Schrödinger functionals related to timely and difficult nonlinear problems in Quantum Mechanics and Nonlinear Optics. We show that these gradients act as preconditioners over traditional choices of descent directions in minimization methods and show a computationally inexpensive way to obtain them using a discrete Fourier basis and a Fast Fourier Transform. We show that the Sobolev preconditioning provides a great convergence improvement over traditional techniques for finding solutions with minimal energy as well as stationary states and suggest a generalization of the method using arbitrary linear operators.

Key words. Sobolev gradients, ground states, Nonlinear Schrödinger equations.

AMS subject classifications. 65K10, 35Q55, 78M50, 82D50.

1. Introduction. The observation of Nature reveals that many unforced continuous systems tend to accommodate into stationary configurations, in which the distributions of mass, charge, velocity, etc, do not change throughout time. In the language of mathematical modeling all configurations are represented by points of a certain space of functions, $\psi(\mathbf{x}) \in W$, while the tendency of the system to lie into any of these states is given by a functional, $E(\psi) : W \to \mathbb{R}$, *the energy*, whose minima are precisely those stationary "states". For this reason it is possible to see many physical problems written as variational principles of the type "find $\psi \in W$ such that $E(\psi) : W \to \mathbb{R}$ achieves a minimum on W". In most situations the functional to be minimized has a dependence on ψ of the form

$$E(\psi) = \int f\left(\nabla\psi(\mathbf{x}), \psi(\mathbf{x})\right) d^n x. \tag{1.1}$$

However, a complete analytical description of the minima of the functional $E(\psi)$ is usually not possible. In this paper we will introduce several techniques for performing this study numerically, focusing on the minimization of $E(\psi)$ subject to physical constraints.

From a practical point of view, this problem is similar to that of finding the minima of a real function defined over a finite-dimensional space, such as $\mathbb{R}^n$. First, a definition of derivative of the functional, $\nabla E(\psi)$, must be chosen. If the domain of the functional W is equipped with some scalar product we may use the Frechet derivative which is given by a first order expansion of the functional around a function ψ

$$E(\psi + \delta) = E(\psi) + \langle \delta, \nabla E(\psi) \rangle + \langle \nabla E(\psi), \delta \rangle + O\left(\|\delta\|^2\right). \tag{1.2}$$

The *critical points,* ψ_c, are defined as the points where the first order variation of the functional vanishes for any perturbation δ. That is, the derivative vanishes in

*Universidad de Castilla-La Mancha, Departamento de Matemáticas, E.T.S.I. Industriales, Avd. Camilo José Cela, 3, Ciudad Real, E-13071, Spain. (jjgarcia@ind-cr.uclm.es)

†Universidad de Castilla-La Mancha, Departamento de Matemáticas, E.T.S.I. Industriales, Avd. Camilo José Cela, 3, Ciudad Real, E-13071, Spain. (vperez@ind-cr.uclm.es).

a weak sense

$$\langle \delta, \nabla E(\psi_c) \rangle = 0, \quad \forall \delta. \tag{1.3}$$

Just like in the finite-dimensional case it is possible to show that any minimum of the functional must be also a critical point. Thus a common approach is to solve Eq. (1.3) and verify *a posteriori* which solutions are actually minima of the functional. If $W = L^2(\mathbb{R}^n)$ this procedure gives us the well known Euler-Lagrange equations of the problem, which is a partial derivatives equation (PDE)

$$\frac{\partial f}{\partial \bar{\psi}} - \nabla \cdot \frac{\partial f}{\partial \nabla \bar{\psi}} = 0. \tag{1.4}$$

However, we have no guarantee to reduce the complexity of the problem, as it is by no means trivial to solve Eq. (1.4). Also we are likely to obtain more solutions than we actually need, since not only minima, but also maxima and saddle node points will satisfy the Lagrange equations.

To avoid these problems some other methods are used which aim at finding the minima of the functional directly, constructing minimizing sequences, $\{\psi_i\}$, whose limit is a minimum of the functional: $\psi_c = \lim_{i\to\infty} \psi_i$. These methods will be discussed in the following sections.

The outline of this paper is as follows. In §2 we recall the definition of Sobolev gradients as given in [2, 3]. We derive a formal solution to the problem of finding these gradients which is based on the inversion of a positive hermitian operator. In §3 we derive an explicit expression for the Sobolev gradients in the trigonometric Fourier basis and comment on its implementation using Fast Fourier Transforms (FFT).

In §4 and §5 we apply the previous tools to two physical problems. Using descent techniques with Sobolev gradients over Fourier spaces we will find the ground states of a Bose-Einstein condensate in a rotating magnetic trap and the excited states for coupled laser beams propagating through a nonlinear medium. Both physical systems are modeled by nonlinear equations of Schrödinger type and present difficulties when traditional minimization techniques are used. We comment on the great improvements that are achieved using Sobolev gradients. Finally in §6 we summarize our results and offer some conclusions.

2. Sobolev gradients.

2.1. Direct solutions of the variational problem. There are two traditional approaches to the problem of finding the minima of a functional using a discrete basis. The first one expands the unknown solution using a Fourier basis $\{\phi_k\}$, $\psi = \sum_k c_k \phi_k$, and defines a new functional over the finite–dimensional space

$$E(\{c_k\}) \equiv E\left(\sum c_k \phi_k\right). \tag{2.1}$$

The functional is then minimized using methods which are well known from the the domain of finite–dimensional problems, e.g. Newton's method or nonlinear conjugate gradient.

This procedure is quite straightforward and there is a huge amount of literature and tools which can be immediately applied to (2.1). However, for some types of problems one has to deal with highly nonlinear algebraic equations with many of terms on each equation, something which is computationally too expensive to work with.

The second approach involves what is known as *descent techniques*. The idea is to manipulate the original functional (1.1) building an analytic equation for a minimizing trajectory in the target space W. This equation is then discretized and solved on a suitable basis. In the *continuous steepest descent* version, the trajectory $\psi(t) : \mathbb{R} \to W$ is continuous and defined by a PDE which involves the gradient of the functional (1.2)

$$\frac{\partial \psi}{\partial t} = -\nabla E. \tag{2.2}$$

The *discrete steepest descent* technique is computationally cheaper since instead of requiring an integrator for the PDE it constructs a discrete succession of estimates to the minimum, $\{\psi_{k+1} = \psi_k + \lambda_k \nabla E(\psi_k)\}$, by locally minimizing $E(\psi_k + \lambda \nabla E)$ with respect to the real parameter λ.

In this paper we will only deal with descent techniques. The first and most important reason is that we will work with the definition of $\nabla E(\psi_k)$ trying to improve its convergence. As it was already shown in [3], this work pays off: a good choice of the gradient improves convergence by several orders of magnitude. The second motivation is that by focusing on the gradient our algorithms will be essentially independent on the descent method, which leaves space for further improvement. For instance one might apply these techniques to a nonlinear conjugate gradient method —which dynamically adjusts the search direction, $d_k \equiv \nabla E(\psi_k)$, with an estimate that takes into account the history of the evolution—. Finally, by working with Eq. (2.2) or its discrete version we avoid the complex nonlinearities that arise in other methods, and we will have a more ample choice of Fourier basis to work with.

2.2. Ordinary gradients. It is customary in the literature to work in spaces which are equipped with an L^2 scalar product and its corresponding norm

$$\langle \psi, \phi \rangle_{L^2} \equiv \int \bar{\psi} \phi, \tag{2.3}$$

$$\|\psi\|^2_{L^2} \equiv \int |\psi|^2. \tag{2.4}$$

If one does so and works with Eq. (1.2) then the formal definition of the gradient is Lagrange's one

$$\nabla E(\psi) = \frac{\partial E}{\partial \bar{\psi}} - \nabla \frac{\partial E}{\partial \left(\nabla \bar{\psi}\right)}. \tag{2.5}$$

We will refer to this definition as the "ordinary" gradient, to distinguish it from the different definitions that we will derive below.

2.3. Sobolev gradients. Following the ideas from [3] we will move our problem to a different space which is the Sobolev space of functions, such that ψ and its derivatives, $\nabla\psi$, have a well defined L^2-norm:

$$\mathbb{H}^1 \equiv \{\psi / \psi, \nabla\psi \in L^2\}. \tag{2.6}$$

This Sobolev space will also be equiped with a scalar product and a norm

$$\langle \psi, \phi \rangle \equiv \int \left[\bar{\psi}(\mathbf{x})\phi(\mathbf{x}) + \nabla\bar{\psi}(\mathbf{x}) \cdot \nabla\phi(\mathbf{x})\right] d^n x, \tag{2.7}$$

$$\|\psi\|^2 \equiv \int \left[|\psi|^2 + |\nabla\psi|^2\right] d^n x. \tag{2.8}$$

To obtain a new explicit expression for the gradient of the functional in the Sobolev space we will follow a less rigorous derivation than in [2]. Performing a first order expansion of our functional around a trial state ψ we obtain

$$E(\psi+\varepsilon\delta) = E(\psi) + \varepsilon \int \overline{(\delta, \nabla\delta)} \begin{pmatrix} \frac{\partial E}{\partial \bar{\psi}} \\ \frac{\partial E}{\partial \nabla \bar{\psi}} \end{pmatrix} + c.c. + O(\varepsilon^2). \tag{2.9}$$

We have to turn this expression into something like Eq. (1.2). This means that we have to find some ϕ such that

$$\int \left[\bar{\delta} \frac{\partial E}{\partial \bar{\psi}} + \nabla \bar{\delta} \frac{\partial E}{\partial (\nabla \bar{\psi})} \right] + c.c = \int \left[\bar{\delta}\phi + \nabla\bar{\delta}\nabla\phi \right] + c.c. \tag{2.10}$$

If we integrate by parts and impose that this equality be satisfied for all perturbations, δ, the problem has a formal solution which is given by a Lagrange equation

$$(1-\triangle)\,\phi = \frac{\partial E}{\partial \bar{\psi}} - \nabla \frac{\partial E}{\partial (\nabla \bar{\psi})}. \tag{2.11}$$

In consequence, our formal expression for the Sobolev gradient of $E(\psi)$ finally reads

$$\nabla_S E \equiv (1-\triangle)^{-1} \nabla E. \tag{2.12}$$

Here $\nabla_S E$ stands for the Sobolev gradient, ∇E is the ordinary one, and $(1-\triangle)^{-1}$ represents the inverse of a linear and strictly positive definite operator.

3. Sobolev gradients on Discrete Fourier spaces. In the rest of the paper we will work with functions which are defined over a d-dimensional rectangular volume $\Omega \equiv \{x \in \Pi_i[a_i, b_i]\}$ with side lengths given by $L_i = b_i - a_i$. We customarily define an orthogonal set of basis functions, $\phi_n = e^{ik_n x}$, over Ω where $\left\{ k_n = 2\pi \left(\frac{n_1}{L_1}, \ldots, \frac{n_d}{L_d} \right),\, n_i \in Z \right\}$

It is well-known that it is possible to expand any continuous function $f(x)$ with periodic boundary conditions using this basis

$$f(x) = \sum_{n=-\infty}^{+\infty} \hat{f}_n \phi_n(x) \tag{3.1}$$

where

$$\hat{f}_n = \frac{1}{V} \int_\Omega \bar{\phi}_n(x) f(x). \tag{3.2}$$

In the previous formula $V = \Pi_i L_i$ is the volume of Ω and arises because of the lack of normalization of the basis functions, a common practice which saves some computation time.

To discretized the problem we will work within the set of functions sampled over a set of evenly spaced points from Ω, $\{x_n = (n_1 h_1, \ldots, n_d h_d),\, n_i = 0, \ldots, N_i - 1\}$. Here h_i represents the spacing along the i-th dimension, n is a vector of non–negative integers and we denote a sampled function by an index, as in $f_n \equiv f(x_n)$.

Due to this discretization our previous Fourier basis is now redundant. It can be reduced to a finite subset of functions which may represent any sampled function.

These functions are given by $\left\{k_n = 2\pi\left(\frac{n_1}{L_1},\ldots,\frac{n_d}{L_d}\right),\ n_i = -M_i+1,\ldots,M_i\right\}$, where $M_i = [N_i/2]$ is the integer quotient of N_i divided by two. In this basis a sampled function is given by an expansion which reads

$$f_m = \sum_n \hat{f}_n \phi_n(x_m), \tag{3.3}$$

and which makes use of the same coefficients as Eq. (3.2).

The advantage of the finite Fourier basis over other approaches is that it provides an approximant of any function whose error is of order $O(L_i/N_i)^p$ where p is the maximum differentiability of the sampled function. Furthermore, there is a numerically efficient method known as Fast Fourier Transform which allows one to compute the Fourier coefficients up from the sampled function, $f_n \to \hat{f}_m$, and vice versa.

Solving Eq. (2.12) numerically in a discrete Fourier basis is simple. Let us say that we have computed the ordinary gradient and that its sampled version has some Fourier coefficients

$$\nabla E(\mathbf{x}_n) = \sum \hat{e}_m \phi_m(x_n). \tag{3.4}$$

and let us assume that there exists a certain solution to Eq. (2.12) and that it has another discrete Fourier expansion

$$\nabla_s E(\mathbf{x}_n) = \sum \hat{s}_m \phi_m(x_n). \tag{3.5}$$

Then by virtue of Eq. (2.12)

$$\hat{s}_m = \frac{\hat{e}_m}{1+k_m^2}. \tag{3.6}$$

That is, in the sampled space the Sobolev gradient represents a preconditioning of the ordinary gradient, such that the most oscillating modes are more attenuated. Furthermore, due to this very simple expression computing the Sobolev preconditioning is computationally cheap and involves only minor changes to existing computer codes based on Fourier transforms.

4. Applications to Quantum Mechanics.

4.1. The problem. In this section we will apply the Sobolev gradients technique to a timely problem from Quantum Physics. The system that we will study is a dilute gas of bosonic atoms which are cooled down to ultralow temperatures at which their dynamics become synchronized. When the temperature is low enough, the gas or "condensate" may be described using a single complex wave function, $\psi(\mathbf{x},t)$ which is ruled by the so called Gross-Pitaevskii equation, a type of Nonlinear Schrödinger equation that for the Bose gas in a rotating trap with adimensional units reads

$$i\partial_t\psi(\mathbf{x},t) = \left[-\frac{1}{2}\triangle + V(\mathbf{x}) + g|\psi(\mathbf{x},t)|^2 - \Omega L_z\right]\psi(\mathbf{x},t). \tag{4.1}$$

Here $g \in \mathbb{R}^+, \Omega \in \mathbb{R}$, $L_z = i\left(x_1\partial_2 - x_2\partial_1\right)$ is an hermitian operator whose expected value $\langle L_z\rangle = \int \bar{\psi} L_z \psi$, represents the *angular momentum* of the condensate along an axis of the trap.

There is conserved quantity associated to equation (4.1) which is called the *energy functional* of the condensate

$$E(\psi) = \frac{1}{2}\int \left\{ |\nabla\psi|^2 + \bar{\psi}\left[V(\mathbf{x}) + \frac{1}{2}g|\psi|^2 - \Omega L_z \right]\psi \right\} d^n x. \tag{4.2}$$

Our objective in this part of the work will be to *find the solutions which are the minima of the energy subject to a restriction of the L^2-norm*

$$\int |\psi|^2 \equiv N. \tag{4.3}$$

The particular value of N is imposed by the experimental conditions. Furthermore, without this restriction the absolute minimum of the energy is always reached at the trivial solution $\psi = 0$.

In Quantum Mechanics the variational formulation of the problem is traditionally converted into a Lagrange equation (1.4), which is nothing but the Gross-Pitaveskii equation for so called *stationary states*. In short the word *"stationary"* refers to solutions of the type

$$\psi(\mathbf{x}, t) = e^{-i\mu t}\phi_\mu(\mathbf{x}). \tag{4.4}$$

The pair $\{\mu, \phi_\mu(\mathbf{x})\}$ satisfies a nonlinear eigenvalue problem

$$\mu\phi_\mu(\mathbf{x}) = \left[-\frac{1}{2}\triangle + V(\mathbf{x}) + g|\phi_\mu(\mathbf{x})|^2 - \Omega L_z \right]\phi_\mu(\mathbf{x}). \tag{4.5}$$

Due to the difficulty of solving directly the problem (4.5) we will try to find a direct solution to the variational problem.

4.2. Numerical methods: imaginary time evolution. We can search the minima of (4.2) using descent techniques modified to account for the restriction on the norm (4.3). The first way to do this is to use a version of the continuous steepest descent which is known as *imaginary time evolution* [**?**]. We can summarize this method with the following set of equations

$$\nu(\mathbf{x}, \tau) = \sqrt{\frac{N}{\|\sigma\|^2_{L^2}}}\sigma(\mathbf{x}, \tau), \tag{4.6}$$

$$\frac{\partial\sigma}{\partial\tau}(\mathbf{x}, \tau) = -\nabla E(\nu), \tag{4.7}$$

$$\nabla E(\nu) = \left[-\frac{1}{2}\triangle + V(\mathbf{x}) + g|\nu|^2 - \Omega L_z \right]\nu. \tag{4.8}$$

Here we see that $\nu(\mathbf{x}, \tau)$ evolves continuously maintaining a fixed L^2–norm N and following the direction of decreasing energy given by ∇E. Indeed it is easy to show that $\frac{\partial}{\partial\tau}\left[E(\nu(\mathbf{x}, \tau))\right] \leq 0$. Hence, the limit given by

$$\phi(\mathbf{x}) = \lim_{\tau\to\infty} \nu(\mathbf{x}, \tau) \tag{4.9}$$

is at least a critical point of the energy, if not a minimum[1]

[1] As it is common with these local minimization procedures, it remains the problem that iterations may be trapped in a critical point which is not a minimum. Linear stability analysis may then be applied to check the validity of the solution.

From a practical point of view, the simplest way to find the minimizer using imaginary time evolution is to repeatedly integrate Eq. (4.7) for a very short time, Δt, apply Eq. (4.6) for the newly found $\sigma(\mathbf{x}, t+\Delta t)$ and use the new estimate for ν to redefine $\sigma \equiv \nu$ and repeat the procedure until convergence. This way one avoids the problem that according to Eq.(4.7) the norm of σ may grow indefinitely. The same consideration applies to the remaining methods that we will present here.

Although the method from Eq. (4.7) was derived using an ordinary gradient, nothing prevents us from applying our Sobolev preconditioning and all results should be still valid. If we do so, our new equations are

$$\nu(\mathbf{x},\tau) = \sqrt{\frac{N}{\|\sigma\|^2_{L^2}}}\sigma(\mathbf{x},\tau), \tag{4.10}$$

$$\frac{\partial\sigma}{\partial\tau}(\mathbf{x},\tau) = -\nabla_S E(\nu), \tag{4.11}$$

$$\nabla_S E(\nu) = (1-\triangle)^{-1}\left[-\frac{1}{2}\triangle + V(\mathbf{x}) + g|\nu|^2 - \Omega L_z\right]\nu, \tag{4.12}$$

and the critical point is still given by Eq. (4.9).

4.3. Numerical methods: minimization of the free energy. While the imaginary time evolution is easy to understand and to implement, the fact that it performs the descent over a path of functions with a certain norm makes it too restricted and sometimes too slow. A different approach is to define a new functional called the *free energy* with a Lagrange multiplier that takes care of the restriction on the norm. In Quantum Mechanics this free energy is usually defined as

$$F_{QM}(\psi) = E(\psi) - \mu N(\psi), \tag{4.13}$$

because it preserves the linear form of the equations. Since our equations are already nonlinear we define a free energy functional more conveniently as

$$F(\psi) = E(\psi) + \frac{1}{2}\left(N(\psi) - \lambda\right)^2. \tag{4.14}$$

First and most important it is not difficult to show that any absolute or relative minimum of $F(\psi)$ is also a minimum of $E(\psi)$ subject to Eq. (4.3) with a nonlinear eigenvalue given by $\mu = N(\psi) - \lambda$.

Secondly, as we will prove in Appendix A, $F(\psi)$ must have at least one finite norm minimum, something which cannot be easily assured for F_{QM}.

The practical advantage of our new functional consists in that fixing λ and Ω we can perform a continuous descent over the whole domain of $F(\psi)$ without renormalizing the solution on each iteration —i.e., the search space is larger. The new equation that we must integrate is thus

$$\frac{\partial\nu}{\partial\tau}(\mathbf{x},\tau) = -\nabla F(\nu) = -\left[-\frac{1}{2}\triangle + V(\mathbf{x}) + g|\nu|^2 - \Omega L_z\right]\nu. \tag{4.15}$$

Using a Sobolev preconditioning Eq. (4.15) becomes

$$\frac{\partial\nu}{\partial\tau}(\mathbf{x},\tau) = -\nabla_S F(\nu) = -(1-\triangle)^{-1}\left[-\frac{1}{2}\triangle + V(\mathbf{x}) + g|\nu|^2 - \Omega L_z\right]\nu. \tag{4.16}$$

In both cases the actual solution is still given by the limit of Eq. (4.9).

4.4. Numerical results. Up to this point we have shown four different numerical methods, two of them incorporating a Sobolev preconditioning [Eqs. (4.11) and (4.16)] and two without it [Eqs. (4.7) and (4.15)]. We have compared the efficiencies of these methods for several test situations. The details of our study are as follows.

All methods have been implemented using the discrete Fourier basis mentioned above. This applies both to the calculation of derivatives and to the application of the Sobolev preconditioning. We restricted our simulations to two-dimensional problems with radially symmetric potential, $V(\mathbf{x}) = \frac{1}{2}|\mathbf{x}|^2$, over a grid of 128×128 points. Practical experience with more complex problems shows that our results generalize to higher dimensionality and denser grids.

Due to the requirements of the method, both variants of imaginary time evolution [Eq. (4.7) and (4.11)] are integrated using a Runge-Kutta-Fehlberg, where the tolerance is adapted as the solution converges to its target. On the other hand, instead of performing a continuous descent for both variants of the free energy descent [Eq. (4.15) and (4.16)], we found it more convenient and faster to perform a discrete steepest descent.

All programs have been implemented using the tensor-algebra environment called Yorick [6], an interpreted environment which is capable of fast numerical computations and which can be equipped with the FFTW library [7]. Execution times are given for a Digital Personal workstation 500au, but the programs run equally well on modest personal computers with Pentium-II processors and less than 64 Mb of memory.

As test case we have considered three situations. Case A is the simplest one and corresponds to a stationary trap $\Omega = 0$, an intense nonlinearity $g = 100$ and starting the minimization with a radially symmetric Gaussian of unit width, that is, $\psi_0 \propto e^{-|\mathbf{x}|^2}$, a shape which is similar to that of the true ground state.

Cases B and C involve a rotating condensate with $\Omega = 0.6$ and $g = 100$. Under these conditions the functional achieves both an absolute minimum, which is the same that found in case A, and a local minimum. The local minimum is a solution of vortex type, a topological defect, whose behavior near zero is $\psi \propto (x_1 + ix_2)/|\mathbf{x}|^2$.

For this reason we designed two test cases with different initial conditions. Case B starts with a Gaussian profile with a centered vortex or $\psi_0 \propto |\mathbf{x}|e^{-|\mathbf{x}|^2}(x_1 + ix_2)/|\mathbf{x}|^2$. In this case all methods are trapped on the local minimum with the centered vortex.

Finally, a third set of simulations, case C, uses the same parameters $\{\Omega = 0.6, g = 100\}$ but starts from a nonsymmetric initial state, $\psi_0 \propto |\mathbf{x}|e^{-|\mathbf{x}|^2}((x_1 - y_1) + i(x_2 - y_1))/|\mathbf{x} - \mathbf{y}|^2$ which is close to the vortex solution but belongs to the basin of attraction of the ground state. Here all minimization methods require more computational work since they must find a path out of the local minimum, and it is precisely in this case where the differences between methods are best shown.

In Table 4.1 we summarize the results of the simulations. In case A it is apparent that the Sobolev preconditioning has a positive influence over convergence, with an astonishing result of 55 steps for the steepest descent with free energy. A similar behavior is found in case B.

In case C, the Sobolev preconditioning enhances convergence speed by two orders of magnitude. An intuitive explanation of why the steepest descent with a Sobolev gradient takes less steps to converge will be discussed in detail in Appendix B.

5. Applications to Nonlinear Optics.

5.1. The model. In this section we will consider a model for a pair of incoherently interacting light beams. To be precise we will study the light field of each

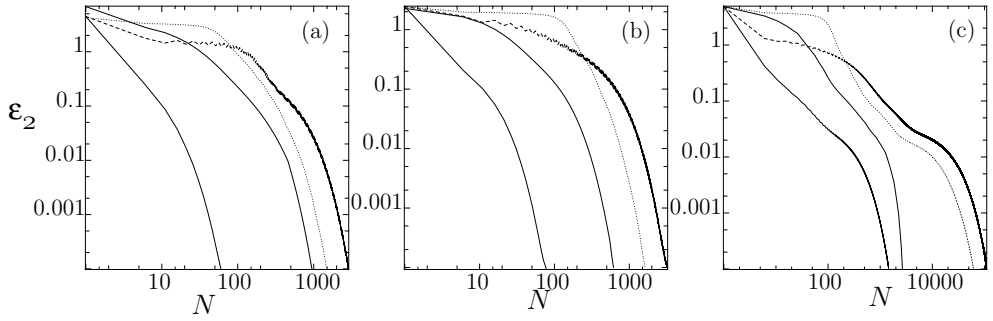

FIG. 4.1. *Evolution of error, $\varepsilon_2 \equiv \|\psi - \psi_{exact}\|_2$, through different minimization processes for continuous steepest descent with Sobolev preconditioning (lower solid line) and without it (dashed line), and for imaginary time evolution with Sobolev preconditioning (upper solid line) and without it (dotted line). Plots (a) to (c) correspond respectively to the cases A,B, and C described in the text. Both axes, error and number of iterations, are in a logarithmic scale.*

Methods		IT	ITS	FE	FES
Case A	Iterations	1320	945	2850	55
	Time (s)	416	371	285	13
Case B	Iterations	1630	615	3210	320
	Time (s)	468	242	75	9
Case C	Iterations	64195	2665	108505	1455
	Time (s)	19863	1165	10861	168

TABLE 4.1

Iterations and computation time for each minimization method: imaginary time without (IT) and with (ITS) Sobolev preconditioning and free energy without (FE) and with (FES) Sobolev preconditioning. Shown are results for the initial data described in the text (cases A, B and C).

beam, $u(\mathbf{x},t)$ and $w(\mathbf{x},t)$, propagating through a weakly nonlinear saturable optical medium. This system may be modeled by the Cauchy problem

$$i\partial_t u = -\triangle u + \frac{u}{1+\kappa(|u|^2+|w|^2)}, \tag{5.1}$$

$$i\partial_t w = -\triangle w + \frac{w}{1+\kappa(|u|^2+|w|^2)}. \tag{5.2}$$

for the complex functions $u, w : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{C}$, which vanish at infinity and satisfy the initial data $u(\mathbf{x},0) = u_0(\mathbf{x})$ and $w(\mathbf{x},0) = w_0(\mathbf{x})$. Here $\kappa \in \mathbb{R}^+$, $-\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator which accounts for the diffraction of light and the nonlinear term $(1+|u|^2+|w|^2)^{-1}$ models the saturable interaction among the beams.

Let us define the two component vector

$$\tilde{U}(\mathbf{x},t) = \begin{pmatrix} u(\mathbf{x},t) \\ w(\mathbf{x},t) \end{pmatrix}, \tag{5.3}$$

then, the energy functional for the system reads

$$E(\tilde{U}) = \int \left[-\tilde{U}^\dagger \triangle \tilde{U} + G(|\tilde{U}|^2) \right] d^n x. \tag{5.4}$$

where $G(\rho) = \frac{1}{\kappa^2}\left(\ln(1+\kappa\rho) - \kappa\rho\right)$. The analysis of this section may be generalized to more general nonlinearities with only minor changes to $G(\rho)$.

5.2. Stationary solutions. We are interested on stationary solutions, which are of the form

$$\tilde{U}(\mathbf{x},t) = \begin{pmatrix} e^{i\mu_u t} & 0 \\ 0 & e^{i\mu_w t} \end{pmatrix} U(\mathbf{x}) = e^{iMt}U(\mathbf{x}). \tag{5.5}$$

The equations for the stationary solutions now are of elliptic type

$$MU = -\triangle U + G'(|U|^2)U. \tag{5.6}$$

with zero Dirichlet boundary conditions at infinity. Again this formulation poses a nonlinear eigenvalue problem for the pair $\{M, U\}$.

The stationary solutions are critical points of the energy functional subject to a constraint on the L^2-norm of each component. That is, defining

$$N_u = \int |u|^2 d^n x, \tag{5.7}$$

$$N_w = \int |w|^2 d^n x, \tag{5.8}$$

the first order variation of the energy around U_0 for fixed N_u and N_w must be zero

$$\left.\frac{\delta E}{\delta U}\right|_{N_u, N_w} = 0. \tag{5.9}$$

The particular fixed values of $\{N_u, N_w\}$ represent the total intensity of each beam of light.

5.3. Ground states. In principle there are many different stationary solutions of Eq. (5.6). However, if we focus on the ground states (stationary solutions of minimal energy) we can apply some of the methods that we mentioned in §4 with minor modifications to account for the higher dimensionality of the problem. For instance, one may define a free energy functional

$$F(U) = E(U) + \frac{1}{2}\left(N_u - \lambda_u\right)^2 + \frac{1}{2}\left(N_w - \lambda_w\right)^2, \tag{5.10}$$

and minimize it using a discrete steepest descent with Sobolev preconditioning.

By performing this minimization using different parameters $\{\lambda_u, \lambda_w\}$ one obtains nodeless localized solutions for u and w as shown in Fig. 5.1. The precise values of λ_u and λ_w determine the norm of each component.

Let us remark that, up to a global factor, the ground state has the same shape in both envelopes. That is,

$$\begin{aligned} u_0(\mathbf{x}) &= N\sqrt{\xi}\,\rho(|\mathbf{x}|, N), \\ w_0(\mathbf{x}) &= N\sqrt{1-\xi}\,\rho(|\mathbf{x}|, N), \quad \forall \xi \in [0,1]. \end{aligned} \tag{5.11}$$

The common shape ρ depends on the total intensity $N = N_u + N_w$, and it is the one that fixes the values of μ_u and μ_w. For this reason if we had chosen the traditional definition of the free energy $F_{QM}(U) = E(U) + \mu_u N_u + \mu_w N_w$ we would have found an infinite degeneracy which prevents convergence to the desired values of $\{N_u, N_w\}$. This problem is removed by the use of our nonlinear Lagrange multipliers and the $\{\lambda_u, \lambda_w\}$ parameters.

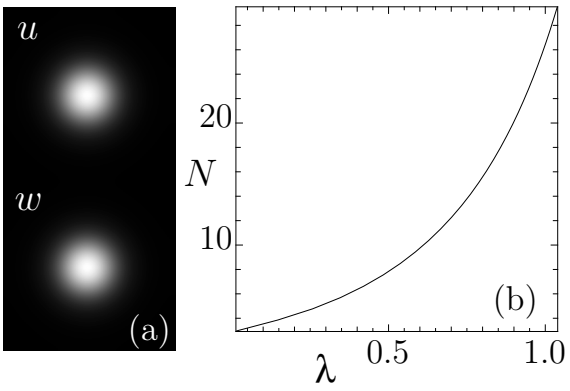

FIG. 5.1. *(a) Density profile of a ground state for $N_u = N_w = 30$. (b) Norms $N_u = N_w = N$, as a function of the nonlinear Lagrange multipliers $\lambda_u = \lambda_w = \lambda$.*

5.4. Excited states. In the field of guided light waves there has been a great interest on the properties of solutions of Eq. (5.6) which are not ground states, the so called *excited states*. Minimization methods based on the energy functional [See §4] cannot be applied to this task since excited states are not necessarily local minima of the energy. Instead, a variational principle somehow equivalent to Eq. (5.6) must be defined.

Let us rewrite Eq. (5.6) as the application of a nonlinear operator

$$f(U) \equiv \left[-\triangle - M + I(|U|^2)\right] U = 0, \tag{5.12}$$

define the error functional

$$F(U) \equiv \int f(U)^\dagger f(U) \geq 0 \tag{5.13}$$

and take our variational principle to be *"find U_0 such that F reaches an absolute minimum $F(U_0) = 0$."*

With this principle and a Sobolev preconditioning our descent technique becomes

$$\frac{\partial \nu(\mathbf{x},\tau)}{\partial \tau} = -\nabla_S F(\nu) = -\left(1 - \triangle\right)^{-1} \nabla F, \tag{5.14}$$

and it may be proven that $U_0(\mathbf{x}) = \lim_{\tau\to\infty} \nu(\mathbf{x},\tau)$. The ordinary gradient ∇F reads

$$G \equiv (-\triangle + I(|U|^2))U, \tag{5.15}$$

$$\nabla F = (-\triangle + I(|U|^2))G + I'(|U|^2)\left(U^\dagger G + G^\dagger U\right) U, \tag{5.16}$$

Its calculation using a discrete Fourier basis is straightforward.

There are several advantages of this approach. The first one is that $F(U)$ makes no distinction between ground and excited states: any stationary state with the right eigenvalues $\mu_i = M_{ii}$ is a local minimum of this new functional. The second one is that we do not need to add any Lagrange multipliers to $F(U)$ since they are already present in the M operators. Finally we expect that the number of minima of $F(U)$ be discrete and separated so as to avoid problems with descent methods being trapped on critical points that are not minima. Indeed, it is very easy to know when this

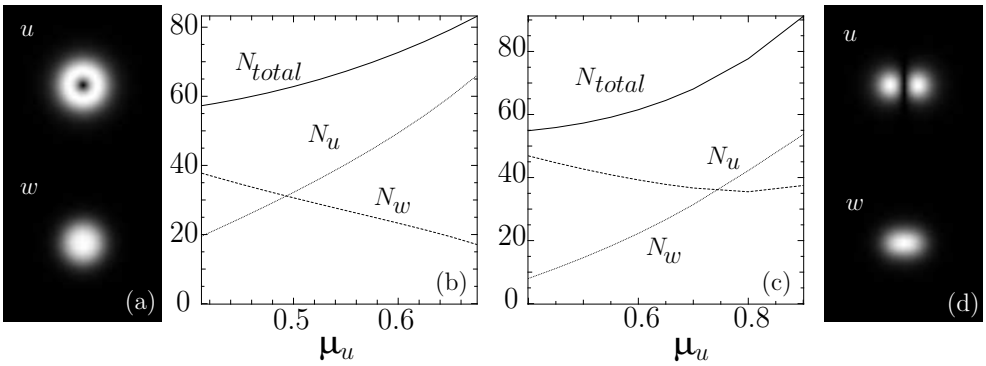

FIG. 5.2. *(a) Vortex-mode vector solitons and (b) their norms N_u, N_w, $N_{total} = N_u + N_w$, as a function of μ_u for $\mu_w = 1$. (d) Dipole-mode vector solitons and (c) their norms N_u, N_w, $N_{total} = N_u + N_w$, as a function of μ_u for $\mu_w = 1$.*

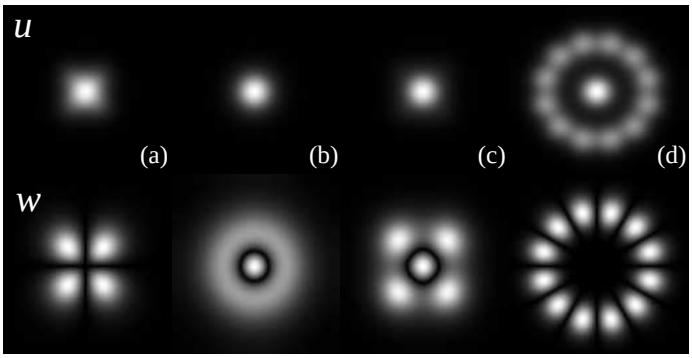

FIG. 5.3. *Different unstable multi–solitonic configurations arising from Eq. (5.14) with a change of the initial conditions for fixed $\mu_u = -1, \mu_w = -0.3, \kappa = 0.5$.*

accidental trapping happens since any absolute minimum of F must also be a zero of it $F(U_0) = 0$.

A remarkable feature of the method that we have outlined above is that to distinguish among the different excited states we have to change both the ad-hoc eigenvalues, $\{\mu_u, \mu_w\}$, and the initial data of the minimization method.

We have concentrated on two types of *excited states* of particular interest for applications: the first one is called *vortex vector soliton* [8], its features being summarized in Fig. 5.2(a-b). Choosing different initial data for the minimization process we obtain an second type of excited states which are called *dipole-mode vector solitons* [9]. An example of these asymmetric solutions is depicted in Fig. 5.2(d). Depending on the parameters $\mu_{u,w}$ we find different norms of the solution. The fact that our method allows the finding of these nonsymmetric stationary states is very interesting since conventional approaches to the problem have severe difficulties. In fact, the application of the Sobolev preconditioning not only greatly enhances the convergence but it is necessary to get convergence.

Nevertheless the method works equally well for more complicated stationary solutions. In Fig. 5.3 we show several exotic solutions that are obtained by solving Eq. (5.14) with different initial conditions. All of those states are dynamically unstable and could have not been found with traditional minimization methods.

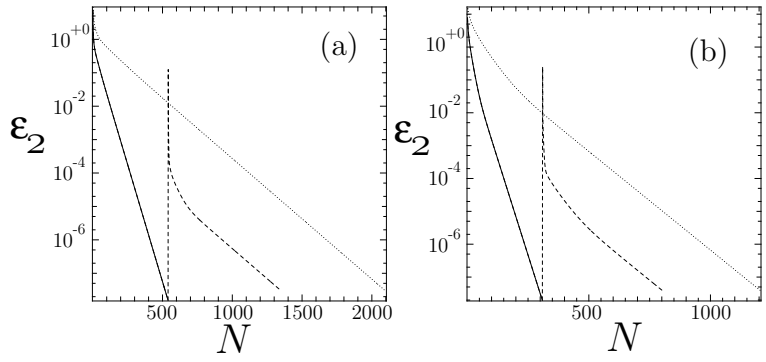

FIG. 5.4. *Dependence of the L^2 error on the number of iterations when looking for a (a)* vortex-mode vector soliton *or a (b)* dipole-mode vector soliton. *We show results for a grid with 32×32 points (solid), 64×64 points (dotted) and 64×64 points starting from an interpolation of the 32×32 solution (dashed).*

5.5. Performance and grid refinement. In this subsection we want to analyze the performance of the method from §5.4 and to introduce a multigrid–like technique to improve convergence rates while looking for more accurate solutions.

The idea of the multigrid technique is to use solutions from coarse-grain grids to calculate better approximations on finer grids [10]. Roughly the algorithm consists in setting a coarse-grain initial data, solving the equation or the variational principle with this initial data until the error is small enough, interpolating the result over a finer grid and iterating using this new grid until both the error and the spatial discretization of the solution are the ones we desire.

Since we are already working with Fourier modes over discrete grids the logical choice for our algorithm is indeed Fourier interpolation. The idea is to use the expansion from Eq. (3.3) *outside of the original grid*, that is,

$$\psi^{(new)}(x) \equiv \sum_n c_n e^{i\mathbf{k}_n \mathbf{x}}, \quad \forall \mathbf{x} \in \Omega. \tag{5.17}$$

This approximation is then discretized over a finer grid, resulting an expansion with a larger number of modes

$$\psi_m^{(new)} = \sum_n c^{(new)} e^{i\mathbf{k}_\mathbf{n}^{(\mathbf{new})} \mathbf{x}_m}, \tag{5.18}$$

which may be used as the starting point for further iteration.

In our case we have used only two different grids, one with 32×32 points (*coarse grid*) and a other with 64×64 points (the *fine grid*). We have used two different initial data for our minimization method: either a pair of Hermite modes which resemble the desired shape, or the solution of the coarse grid interpolated over the finer grid. As both the evolution of the error in Fig. 5.4 and the computation times in Table 5.1 show, interpolation saves a significant amount of time. Indeed, the interpolated solution with only 32×32 Fourier modes is already a good approximation for the fine grid, as it shows the small change of the error in Fig. 5.4

6. Conclusions. In this paper we have shown a numerically efficient way to improve the convergence of several minimization methods using the so called Sobolev gradients and applied it to different problems which involve Nonlinear Schrödinger

	Initial data type					
	Hermite modes for vortex solitons.		Interpolated solution for vortex	Hermite modes for dipole solitons.		Interpolated solution for dipole
Grid	32×32	64×64	64×64	32×32	64×64	64×64
Iterations	540	2101	812	496	1206	496
Time (s)	106	2006	757	60	1141	466

TABLE 5.1

Results for the search of stationary solutions of Eq. (5.1)-(5.2), with and without grid refinement. Shown are the results for different initial data. The initial data named "interpolated solution" corresponds to taking as initial data the approximated solution on the 32×32 *grid.*

Equations. We have also proven that these gradients represent a preconditioning over the traditional definition of gradients on L^2. In Appendix B we suggest a generalization of this method to different vector spaces and scalar products.

We have presented two different methods for solving our minimization problems: a traditional one, the *imaginary time evolution*, and a new one, *the minimization of a nonlinear free energy*. Both methods have been shown to be suitable for a Sobolev preconditioning, giving us two new methods that we call *preconditioned imaginary time evolution* and a *preconditioned free energy*.

We have implemented all four methods using a discrete Fourier basis and a Fast Fourier Transform. With these tools we have shown that the Sobolev preconditioning becomes an inexpensive additional step over existing methods. The four resulting solvers have been applied to several realistic problems and in all tests the nonlinear free energy with the Sobolev preconditioning showed the best convergence rates. Indeed the Sobolev preconditioning has an important effect on convergence which can be as good as gaining two orders of magnitude over the traditional techniques. Furthermore, opposite to what happens with finite differences [2] the preconditioning may be applied without significant computational cost.

We have derived a new method for finding excited states of coupled nonlinear Schrödinger equations. This method introduces a new variational principle which is not based on an energy functional but on finding the zeros of a nonlinear operator which corresponds to the equation to be solved. We have also shown how to improve convergence using Fourier interpolation and a two-grid method as a source for better initial approximations of the iterative method. This approach may be extended to more sophisticated multigrid methods.

Appendix A. Existence of minimizers for the nonlinear free energy functional..

In this appendix we want to prove the existence of minimizers for the free energy (4.14). Let us write the free energy functional in the following form

$$F(\psi) = \int \bar{\psi} A_\Omega \psi d^n x + \int \frac{g}{2} |\psi|^4 d^n x + \frac{1}{2} \left(N - \lambda\right)^2, \tag{A.1}$$

where $N = \int |\psi|^2 d^n r$ and

$$A_\Omega = -\frac{1}{2}\triangle + V(\mathbf{x}) - \Omega L_z, \quad \Omega \in [0, 1) \tag{A.2}$$

is a positive hermitian operator $A_\Omega \geq \mu_{min} > 0$.

Let us also define the following spaces of functions defined over $\mathbb{R}^n$. We will work in L^p spaces

$$L^p = \left\{ \psi : \mathbb{R}^2 \to \mathbb{C} \ / \ \|\psi\|_p := \left(\int |\psi(\mathbf{x})|^p d^n x \right)^{1/p} < +\infty \right\}. \tag{A.3}$$

Also of interest will be the space of functions over which A_Ω is well defined

$$H_\Omega = \left\{ \psi : \mathbb{R}^2 \to \mathbb{C} \ / \ \|\psi\|_\Omega := \left(\int \bar{\psi}(\mathbf{x}) A_\Omega \psi(\mathbf{x}) \right)^{1/2} < +\infty \right\}. \tag{A.4}$$

Let us remark that the dimensionality of the space, $\mathbb{R}^2$ is important, since for $n \leq 2$ it is easy to show that $H_\Omega \subset L^4$.

With the preceding notation we will state the following relevant theorem:

THEOREM A.1. *The free energy functional F given by Eq. (A.1) has at least one absolute minimum in the set given by the inequalities*

$$0 < \|\psi\|_\Omega^2 \leq \lambda \|\psi\|_2^2 \leq \lambda(\lambda - \mu_{min}).$$

Proof. The first step of the proof will be to show that the domain of F is indeed the whole space H_Ω. Using the positivity of the A_Ω operator we show that

$$\|\psi\|_2 \leq \sqrt{\mu_{min}} \|\psi\|_\Omega, \tag{A.5}$$

which means that $H_\Omega \subset L^2$. We need Sobolev's inequality [11]

$$\|\psi\|_k \leq \|\psi\|_2^{1-d} \|\nabla \psi\|_2^d \tag{A.6}$$

where $d = n/2 - n/k$, n is the dimensionality of the space and for us $d = 1/2$. Applying this inequality we obtain the following bound

$$\|\psi\|_4 \leq \sqrt{\|\psi\|_2 \|\nabla \psi\|_2} \leq \mu_{min}^{\frac{1}{4}} \|\psi\|_\Omega, \tag{A.7}$$

which means that $H_\Omega \subset L^4$. Since $\|\cdot\|_2, \ \|\cdot\|_4, \ \|\cdot\|_\Omega < +\infty$ inside H_Ω we conclude that F as given by Eq. (A.1) is well defined over the whole space.

F is also continuous. To prove it let us rewrite the free energy functional as

$$F(\psi) = \|\psi\|_\Omega^2 + \frac{g}{2} \|\psi\|_4^4 + \frac{1}{2} (\|\psi\|_2^2 - \mu)^2, \tag{A.8}$$

and using the bounds (A.5) and (A.7) it is easy to show that

$$|F(\psi) - F(\xi)| \leq \delta(\varepsilon), \quad \forall \xi \ \|\psi - \xi\|_\Omega \leq \epsilon \tag{A.9}$$

where the constant $\delta(\varepsilon)$ is given by ε and $\|\psi\|_4$.

We will also need to show that F is coercive

$$\lim_{\|\psi\| \to \infty} \frac{F(\psi)}{\|\psi\|} \geq \alpha > 0. \tag{A.10}$$

Using Eq. (A.8) one may show that indeed

$$\frac{F(\psi)}{\|\psi\|_\Omega} \geq \|\psi\|_\Omega, \tag{A.11}$$

and thus the quotient tends to infinity as the norm grows. The continuity and coercitivity of F allow to use theorem 1 (section 1.2) of [12], which states the existence of at least one minimizer $\{\inf F(u) \ : \ u \in X\}$ of F provided it is a weakly lower semicontinuous and coercive application $F : X \to \mathbb{R}$.

So we know that there is at least one minimum, but we do not know where to look for it. Let us now show how the λ parameter allows us to select different targets for the minimization problem. To do so we define a real one-dimensional function for any given direction $\psi \in H_\Omega$

$$f(k) \equiv F(k\psi). \tag{A.12}$$

This function is nothing but a polynomial over k

$$f(k) = kNa_\psi + k^2N^2\frac{u_\psi}{2} + \frac{1}{2}\left(kN - \lambda\right)^2, \tag{A.13}$$

where $a_\psi = \int \bar{\psi}A\psi/N$ and $u_\psi = \frac{g}{2}\int |\psi|^4/N^2$ are constants that depend only on the precise direction ψ.

By differentiating the polynomial and imposing $k = 0$ we find that $f'(0) < 0$ at the origin for all possible directions. This means that, as we mentioned above, our Lagrange "multiplier" λ allows us to avoid the useless solution, $\psi = 0$.

Furthermore, we can restrict the location of the minimizer to a surface of a certain norm. By differentiating $f(k)$ we reach

$$(a_\psi - \lambda) + Nk(u_\psi + 1) = 0. \tag{A.14}$$

This equation has a single solution which gives us the norm of minimal energy along the ψ-direction

$$N_{min}(\psi) = \max\left\{0, \frac{\lambda - a_\psi}{u_\psi + 1}\right\} \leq \lambda - \mu_{min}, \tag{A.15}$$

a value which is bounded above by our Lagrange multiplier λ.

From Eq. (A.14) it also follows that for any absolute minimum of the functional, ψ_{min}, the expected value of the A_Ω operator must be bounded by the L^2 norm

$$\int \bar{\psi}_{min}A\psi_{min} = \|\psi_{min}\|^2_\Omega \leq \lambda N(\psi_{min}). \tag{A.16}$$

Otherwise the trivial solution $\psi = 0$ would have less energy than ψ_{min}. We can thus, instead of working with an unknown surface, delimit the location of the minimum to a set which is given by two inequalities, (A.15) and (A.16)

$$W = \{\psi \in H_\Omega \ : \ 0 < N(\psi) \leq \lambda - \mu_{min}, \|\psi\|^2_\Omega \leq \lambda N(\psi)\}. \tag{A.17}$$

□

There are several practical consequences of this theorem. First, it states that the problem of finding the minima of $E(\psi)$ subject to fixed norm has one solution, i.e.,

there exists at least one ground state. Second, but equally important, it proves our Lagrange penalizer $\frac{1}{2}(N-\mu)^2$ to be specially well suited for this problem since it avoids both the useless solution $\psi = 0$ and those with too large norm. And finally it gives us a bounded in which the minimizer must be. Indeed we can extend this result by proving that we can restrict our search space to a compact superset of W.

THEOREM A.2. *Any absolute minimum, ψ, of the functional F given by Eq. (A.1) lays inside the compact set of L^2*

$$\bar{U} = \{\psi \in L^2 \ : \ \|\psi\|_\Omega^2 \leq \lambda \|\psi\|_2^2 \leq \lambda(\lambda - 1)\}.$$

Proof. For the type of spaces that we work with, a set is compact iff it is closed and we can build an ε-net for any positive number ε. The ε-net is a finite set $K_\varepsilon = \{\nu_1, \ldots, \nu_k\}$ such that for each $\psi \in \bar{U}$ there is an element $\nu \in K_\epsilon$ verifying $\|\psi - \nu\|_\Omega < \varepsilon$. Thus compactness is equivalent to the possibility of building an arbitrarily good approximation of our minimizer using a finite but sufficiently large basis of functions.

It is evident that the set $\bar{U}$ is closed in the subspace H_Ω of L^2. Let $\{u_n\} \in \bar{U}$ be a convergent sequence and let u be their limit. Since for each element of the sequence

$$\|u_n\|_\Omega \leq \sqrt{\lambda}\|u_n\|_2 < \lambda(\lambda - 1), \tag{A.18}$$

it is also obvious that

$$\|u\|_\Omega \leq \sqrt{\lambda}\|u\|_2 \leq \lambda(\lambda - 1), \tag{A.19}$$

also thus the limit belongs to $\bar{U}$.

The compactness of the closed set $\bar{U}$ essentially follows from the fact that the eigenstates of A_Ω form a complete basis of H_Ω, and that the eigenvalues of A_Ω form a monotonously growing unbounded set of positive numbers. These eigenstates are of the form

$$\phi_{n,l} = P_n^l(|\mathbf{x}|)\frac{x_1 + ix_2}{|\mathbf{x}|^2}e^{-|\mathbf{x}|^2/2}, \tag{A.20}$$

where P_n^l are Laguerre's generalized polynomials. And the corresponding eigenvalues are

$$A_\Omega \phi_{n,l} = \mu_{n,l}\phi_{n,l} = (2n + (1-\Omega)l + 1)\phi_{n,l}, \quad n, l \in \mathbb{N} \cup \{0\}. \tag{A.21}$$

Let us choose any natural number k such that $k + 1 > \lambda$. We can split the whole space as a direct sum $H_\Omega = H_0^{k+1} \oplus H_{k+1}^\infty$ where

$$H_j^k = \mathrm{lin}\{\phi_{n,l} \ : \ j \leq 2n + l < k\}. \tag{A.22}$$

The important point is that since H_0^{k+1} is isomorph to $\mathbb{R}^m$ for some natural number m, and $\bar{U}_k = \bar{U} \cap H_0^{k+1}$ lays inside a compact m-dimensional ball of radius $\sqrt{\lambda - 1}$, then we can find any ε-net for $\bar{U}_k$. Furthermore, by separating

$$\psi = \psi_a + \psi_b, \quad \psi_a \in \bar{U}_k, \psi_b \in H_{k+1}^\infty, \tag{A.23}$$

and using the definition of $\bar{U}_k$ we show that the projection of ψ outside of $\bar{U}_k$ can be made arbitrarily small

$$\|\psi_b\|_2^2 \leq \frac{\lambda}{k+1}N. \tag{A.24}$$

A direct consequence of this is that for $k > \lambda N/\varepsilon$, a ε-net of $\bar{U}_k$ is also an ε-net of $\bar{U}$, which proves the compactness.

Finally, by inspecting the eigenvalues of A_Ω and using Eq. (A.15-A.16) it is not difficult to see that the absolute minimum of F must lay in $\bar{U}$.
□

Appendix B. Extending the Sobolev gradients.

We have shown that redefining the gradient turns out to be a kind of preconditioning over the original choice of the direction of descent. Let us assume that our functional has the following form

$$E(\psi) = \int \bar{\psi} A \psi + f(|\psi|^2, \mathbf{x}), \tag{B.1}$$

where A is a non–negative hermitian operator that may involve some derivatives. Let us also assume that H is a suitable space equipped with the following scalar product

$$\langle \psi, \phi \rangle = \int \bar{\psi} \, (1 + A) \, \phi, \tag{B.2}$$

which is indeed a scalar product because $A|_H \geq 0$.

Let us rewrite the energy functional in the following way

$$E(\psi) = \langle \psi, \psi \rangle + \int f(|\psi|^2, \mathbf{x}) - |\psi|^2. \tag{B.3}$$

The Sobolev gradient in H is

$$\nabla_A E = \psi + (1 + A)^{-1} \left[\partial_1 f - \psi \right], \tag{B.4}$$

while the so called ordinary gradient is $\nabla E = A\psi + \partial_1 f$. Hence the preconditioning nature of the method is recovered

$$\nabla_A E = (1 + A)^{-1} \nabla E. \tag{B.5}$$

We can thus think that the new choice of the scalar product aims at making the linear part of the energy functional close to some quadratic form, $\langle \psi, B\psi \rangle$, such that the new operator B is almost the unity. In our example, indeed, $B\psi = \psi$. We that this preconditioning will both enhance the directions of decreasing energy and also have a smoothing effect on the nonlinear part. However, the problem is complicated and we know of no proof that these arguments be of general applicability.

REFERENCES

[1] J. M. Sanz-Serna, *Fourier techniques in numerical methods for evolutionary problems*, in "*Third Granada Lectures on Computational Physics*", Proc. of the III Granada Seminar on Computational Physics of 1994, P. L. Garrido, J. Marro, Springer-Verlag, Berlin, 1995, pp. 145–200.

[2] J. W. Neuberger, R. J. Renka, *Sobolev Gradients and the Ginzburg-Landau functional*, SIAM J. Sci. Comput., 20(1998) pp. 582–590.

[3] J. W. Neuberger, *Sobolev Gradients and Differential Equations*, Springer-Verlag, Berlin, 1997.

[4] J. J. García-Ripoll, V. M. Pérez-García, *Stability of vortices in inhomogeneous Bose–Einstein condensates subject to rotation: a three dimensional analysis*, Phys. Rev. A, 60(1999), pp. 4864-4874.

[5] J. Nocedal, S. J. Wright, *"Numerical Optimization"*, Springer Series in Operations Research, Springer, New York, 1999.

[6] The Yorick environment for scientific computing is available at the electronic address ftp://ftp-icf.llnl.gov/pub/Yorick/yorick-ad.html.
[7] Matteo Frigo, Steven G. Johnson, *FFTW: An Adaptive Software Architecture for the FFT*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, IEEE, Seattle, WA, 1998, pp. 1381–1384. See also
[8] Z. H. Musslimani, M. Segev, D. N. Christodoulides, M. Soljacić, *Composite Multihump Vector Solitons Carrying Topological Charge* Phys. Rev. Lett., 84(1999), pp. 1164–1167.
[9] J. J. García–Ripoll, V. M. Pérez–García, E. A. Ostrovskaya, Y. S. Kivshar, *Dipole–mode vector solitons*, Phys. Rev. Lett., 85(2000), pp. 82–85.
[10] W. L. Briggs, V. E. Henson, S. F. McCormick, "*A multigrid tutorial*", SIAM (2000).
[11] L. Hörmander, "*The analysis of linear partial differential operators I*", Springer, Berlin (1983).
[12] B. Dacorogna, "*Direct Methods in the Calculus of Variations*", Springer–Verlag, Berlin (1989).